PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/04, C07K 14/435, C12Q 1/25, 1/68 — A1

(11) International Publication Number: WO 96/23810

(43) International Publication Date: 8 August 1996 (08.08.96)

(21) International Application Number: PCT/US95/14692

(22) International Filing Date: 13 November 1995 (13.11.95)

(30) Priority Data:
08/337,915 10 November 1994 (10.11.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/337,915 (CIP)
Filed on 10 November 1994 (10.11.94)

(71) Applicant *(for all designated States except US)*: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: TSIEN, Roger, Y. [US/US]; 8535 Nottingham Place, La Jolla, CA 92037 (US). HEIM, Roger [CH/US]; 510 Stratford Court #309B, Del Mar, CA 92014 (US).

(74) Agent: BERLINER, Robert; Robbins, Berliner & Carson, 5th floor, 201 N. Figueroa Street, Los Angeles, CA 90012-2628 (US).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG). European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: MODIFIED GREEN FLUORESCENCT PROTEINS

(57) Abstract

Modifications in the sequence of *Aequorea* wild-type GFP provide products having markedly different excitation and emission spectra from corresponding products from wild-type GFP. In one class of modifications, the product derived from the modified GFP exhibits an alteration in the ratio of two main excitation peaks observed with the product derived from wild-type GFP. In another class, the product derived from the modified GFP fluoresces at a shorter wavelength than the corresponding product from wild-type GFP. In yet another class of modifications, the product derived from the modified GFP exhibits only a single excitation peak and enhanced emission relative to the product derived from wild-type GFP.

MODIFIED GREEN FLUORESCENT PROTEINS

Background of the Invention

This invention relates generally to the fields of biology and chemistry. More particularly, the invention is directed to modified fluorescent proteins and to methods for the preparation and use thereof.

This invention was made with Government support under Grant No. NS27177, awarded by the National Institute of Health. The Government has certain rights in this invention.

In biochemistry, molecular biology and medical diagnostics, it is often desirable to add a fluorescent label to a protein so that the protein can be easily tracked and quantified. The normal procedures for labeling requires that the protein be covalently reacted *in vitro* with fluorescent dyes, then repurified to remove excess dye and any damaged protein. If the labeled protein is to be used inside cells, it usually has to be microinjected; this is a difficult and time-consuming operation that cannot be performed on large numbers of cells. These problems may, however, be [illegible] by joining a nucleotide sequence coding for the protein of interest with the sequence for a naturally fluorescent protein, then expressing the fusion protein.

The green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* is a remarkable protein with strong visible absorbance and fluorescence from a *p*-hydroxybenzylideneimidazolone chromophore, which is generated by cyclization and oxidation of the protein's own Ser-Tyr-Gly sequence at positions 65 to 67. A cDNA sequence [SEQ ID NO:1] for one isotype of GFP has been reported [Prasher, D. C. et al., *Gene* 111, 229-233 (1992)]; cloning of this cDNA has enabled GFP expression in different organisms. The finding that the expressed protein becomes fluorescent in cells from a wide variety of organisms [Chalfie, M. et al., *Science* 263, 802-805 (1994)] makes GFP a powerful new tool in molecular and cell biology and indicates that the oxidative cyclization must be either spontaneous or dependent only on ubiquitous enzymes and reactants.

Summary of the Invention

In accordance with the present invention, it has been determined that particular modifications in the polypeptide sequence of an *Aequorea* wild-type GFP [SEQ ID NO:2] lead to formation of products having markedly different excitation and emission spectra from corresponding products derived from wild-type GFP. Visibly distinct colors and/or increased intensities of emission make these products useful in a wide variety of contexts, such as tracking of differential gene expression and protein localization.

Brief Description of the Drawings

The invention may be better understood with reference to the accompanying drawings, in which:

Fig. 1 compares different versions of GFP by gel electrophoresis and Coomassie blue staining;

Fig. 2 illustrates a proposed biosynthetic scheme for GFP;

Figs. 3a and 3b illustrate the excitation and emission spectra of wild-type and a first group of mutant GFPs;

Figs. 4a and 4b illustrate the excitation and emission spectra of wild-type and a second group of mutant GFPs;

Fig. 5 illustrates the rate of fluorophore formation in the wild-type GFP and the Ser 65→Thr mutant;

Figs. 6a and 6b illustrate the behavior of wild-type GFP and the Ser 65→Thr mutant, respectively, upon progressive irradiation with ultraviolet light; and

Fig. 7 illustrates fluorescence excitation and emission spectra of a third group of GFP mutants.

Detailed Description of the Invention

GFP was expressed in *E. coli* under the control of a T7 promoter for quantitative analysis of the properties of the recombinant protein. Gel electrophoresis under denaturing conditions showed protein of the expected molecular weight (27 kDa) as a dominant band (Fig. 1), which could be quantified

A molecular interpretation is presented in Fig. 2. If the newly translated apoprotein (top left) evades precipitation into inclusion bodies, the amino group of Gly 67 might cyclize onto the carbonyl group of Ser 65 to form an imidazolidin-5-one, where the process would stop (top center) if O_2 is absent. The new N=C double bond would be expected to promote dehydrogenation to form a conjugated chromophore; imidazolidin-5-ones are indeed known to undergo autoxidative formation of double bonds at the 4-position [Kjaer, A. *Acta Chem. Scand.* 7, 1030-1035 (1953); Kidwai, A. R. & Devasia, G. M. *J. Org. Chem.* 27, 4527-4531 (1962)], which is exactly what is necessary to complete the fluorophore (upper right). The protonated and deprotonated species (upper and lower right) may be responsible for the 395 and 470-475 nm excitation peaks, respectively. The excited states of phenols are much more acidic than their ground states, so that emission would come only from a deprotonated species.

The *Aequorea* GFP cDNA was subjected to random mutagenesis by hydroxylamine treatment or polymerase chain reaction. Approximately six thousand bacterial colonies on agar plates were illuminated with alternating 395 and 475 nm excitation and visually screened for altered excitation properties or emission colors.

According to a first aspect of the present invention, modifications are provided which result in a shift in the ratio of the two excitations peaks of the product after oxidation and cyclization relative to the wild type. Three mutants were found with significant alterations in the ratio of the two main excitation peaks (Table I). The mutations were sequenced and recombined with the wild-type gene in different ways to eliminate neutral mutations and assign the fluorescence effects to single amino acid substitutions, except for H9 where two neighboring mutations have not yet been separated. They all lay in the C terminal part of the protein (Table I), remote in primary sequence from the chromophore formed from residues 65-67.

These and other modifications are defined herein with reference to the amino acid sequence [SEQ ID NO:2] encoded by the reported cDNA [SEQ ID NO:1]; the first amino acid identified is the one found at the indicated location in

formation due to steric considerations. Phe gave weak fluorescence with an excitation maximum at 358 nm and an emission maximum at 442 nm. Accordingly, pursuant to this aspect of the invention modified GFP proteins which fluoresce at different wavelengths (preferably, different by at least 10 nm and more preferably, by at least 50 nm) relative to the native protein are provided, for example, those wherein Tyr 66 is replaced by Phe, His or Trp.

In a further embodiment of this aspect of the invention, a double mutant Y66H, Y145F was identified which had almost the same wavelengths as the single mutant Y66H but almost twice the brightness, due mainly to a higher quantum efficiency of fluorescence. The double mutant also developed its fluorescence during overnight growth, whereas the single mutant required several days.

In accordance with further embodiments of this aspect of the invention, a first round of mutagenesis to increase the brightness of Y66W yielded M153T/V163A/N212K as additional substitutions. This mutant was subjected to another round of mutagenesis, resulting in two further sets, N146I and I123V/Y145H/H148R (Table II). The quantum efficiency of these mutants is now comparable to wild-type GFP. The clustering of the substitutions in residues 145 to 163 suggest that those residues lie relatively close to the chromophore and that reductions in the size of their side chains might be compensating for the larger size of tryptophan compared to tyrosine.

Pursuant to yet another aspect of the present invention, modified GFP proteins are provided which provide substantially more intense fluorescence per molecule than the wild type protein. Modifications at Ser 65 to Ala, Leu, Cys, Val, Ile or Thr provide proteins with red-shifted and brighter spectra relative to the native protein. In particular, the Thr mutant (corresponding to a change in the GFP cDNA sequence [SEQ ID NO:1] at 193-195 from TCT to ACT) and Cys mutant (corresponding to a change in the GFP cDNA sequence [SEQ ID NO:1] at 193-195 from TCT to TGT) are about six times brighter than wild type when excited at the preferred long-wavelength band above 450 nm. As a consequence, these modified proteins are superior to wild type proteins for practically all applications. Further, the brightness of these modified proteins matches the

the locations of known neutral mutations suggest that amino acids 76-115 are less critical to the spectroscopic properties of the product. In addition, as would be immediately apparent to those working in the field, the use of various types of fusion sequences which lengthen the resultant protein and serve some functional purpose in the preparation or purification of the protein would also be routine and are contemplated as within the scope of the present invention. For example, it is common practice to add amino acid sequences including a polyhistidine tag to facilitate purification of the product proteins. As such fusions do not significantly alter the salient properties of the molecules comprising same, modified GFPs as described herein including such fusion sequences at either end thereof are also clearly contemplated as within the scope of the present invention.

Similarly, in addition to the specific mutations disclosed herein, it is well understood by those working in the field that in many instances modifications in particular locations in the polypeptide sequence may have no effect upon the properties of the resultant polypeptide. Unlike the specific mutations described in detail herein, other mutations provide polypeptides which have properties essentially or substantially indistinguishable from those of the specific polypeptides disclosed herein. For example, the following substitutions have been found to be neutral (i.e., have no significant impact on the properties of the product): Lys 3→Arg; Asp 76→Gly; Phe 99→Ile; Asn 105→Ser; Glu 115→Val; Thr 225→Ser; and Lys 238→Glu. These equivalent polypeptides (and oligonucleotide sequences encoding these polypeptides) are also regarded as within the scope of the present invention. In general, the polypeptides and oligonucleotide sequences of the present invention (in addition to containing at least one of the specific mutations identified herein) will be at least about 85% homologous, more preferably at least about 90% homologous, and most preferably at least about 95% homologous, to the wild-type GFP described herein. Because of the significant difference in properties observed upon introduction of the specified modifications into a GFP sequence, the presence of the specified modifications relative to the corresponding reported sequence for wild-type GFP [SEQ ID NO:2] are regarded as central to the invention.

expression of both A and B by searching for the presence of both colors simultaneously.

As another example, to examine the precise temporal or spatial relationship between the generation or location of recombinant proteins X and Y within a cell or an organism, one could fuse genes for different colors of GFP to the genes for proteins X and Y, respectively. If desired, DNA sequences encoding flexible oligopeptide spacers could be included to allow the linked domains to function autonomously in a single construct. By examining the appearance of the two distinguishable colors of fluorescence in the very same cells or organisms, one could compare and contrast the generation or location of the proteins X and Y with much greater precision and less biological variability than if one had to compare two separate sets of cells or organisms, each containing just one color of GFP fused to either protein X or Y. Other examples of the usefulness of two colors would be obvious to those skilled in the art.

The further mutations to brighten the Y66H and Y66W variants of GFP enhance the possibility of using two or three colors of fluorescent protein to track differential gene expression, protein localizations or cell fates. For example, mutants P4-3 (Y66H/Y145F), W7 (Y66W/N146I/M153T/V163A/N212K) and S65T can all be distinguished from each other. P4-3 is specifically detected by exciting at 290-370 nm and collecting emission at 420-460 nm. W7 is specifically detected by exciting at 410-457 nm and collecting emission at 465-495 nm. S65T is specifically detected by exciting at 483-493 nm and collecting emission at wavelengths greater than 510 nm. Bacteria carrying these three proteins are readily discriminated under a microscope using the above wavelength bandpass filters.

The chromophore in GFP is well buried inside the rest of the protein, so much of the dimness of the original point mutants was presumably due to steric mismatch between the substituted amino acid and the cavity optimized for tyrosine. The location of the beneficial mutations implies that residues 145-163 are probably close to the chromophore. The M153A/S65T mutant has the longest wavelengths

μM of dGTP, dCTP, and dTTP [Muhlrad, D. et al., *Yeast* 8, 79-82 (1992)]. The product was ligated into pGEMEX2 and subsequently transformed into JM109(DE3). Colonies on agar were visually screened for different emission colors and ratios of brightness when excited at 475 vs. 395 nm.

Figs. 3a and 3b illustrate the excitation and emission spectra of wild-type and mutant GFPs. In Figs. 3a and 3b, —— wild-type; — — S202F,T203I; - - - I167T; ----- Y66W; — • — • Y66H. Samples were soluble fractions from *E. coli* expressing the proteins at high level, except for Y66W, which was obtained in very low yield and measured on intact cells. Autofluorescence was negligible for all spectra except those of Y66W, whose excitation spectrum below 380 nm may be contaminated by autofluorescence. Excitation and emission spectra were measured with 1.8 nm bandwidths and the non-scanning wavelength set to the appropriate peak. Excitation spectra were corrected with a rhodamine B quantum counter, while emission spectra (except for Y66W) were corrected for monochromator and detector efficiencies using manufacturer-supplied correction spectra. All amplitudes have been arbitrarily normalized to a maximum value of 1.0. A comparison of brightness at equal protein concentrations is provided in Table 1.

polymerase using the mutagenic oligonucleotide to prime DNA synthesis; the Version 2 kit utilizes unmodified T7 DNA polymerase to synthesize the complementary strand. When the heteroduplex molecule is transformed into a host with an active uracil-N-glycosylase (which cleaves the bond between the uracil base and the ribose molecule, yielding an apyrimidic site), the uracil-containing wild-type strand is inactivated, resulting in an enrichment of the mutant strand.

The coding region of GFP cDNA was cloned into the *BamHI* site of the phagemid $pRSET_B$ from Invitrogen (San Diego, CA). This construct was introduced into the *dut, ung* double mutant *E. coli* strain CJ236 provided with the Muta-Gene kit and superinfected with helper phage VCSM13 (Stratagene, La Jolla, CA) to produce phagemid particles with single-stranded DNA containing some uracils in place of thymine. The uracil-containing DNA was purified to serve as templates for *in vitro* synthesis of the second strands using the mutagenic nucleotides as primers. The DNA hybrids were transformed into the strain XL1blue (available from Stratagene), which has a functional uracil-N-glycosylase; this enzyme inactivates the parent wild-type DNA strand and selects for mutant clones. DNA of several colonies were isolated and checked for proper mutation by sequencing

To express the mutant proteins, the DNA constructs obtained by mutagenesis were transformed into *E. coli* strain BL21(DE3)LysS (Novagen, Madison, WI), which has a chromosomal copy of T7 polymerase to drive expression from the strong T7 promotor. At room temperature 3 ml cultures were grown to saturation (typically, overnight) without induction. Cells from 1 ml of culture were collected, washed and finally resuspended in 100 μl of 50 mM Tris pH 8.0, 300 mM NaCl. The cells were then lysed by three cycles of freeze/thawing (liquid nitrogen/30° C water bath). The soluble fraction was obtained by pelletting cell debris and unbroken cells in a microfuge.

To facilitate purification of the recombinant proteins, the vector used fuses a histidine tag (6 consecutive His) to the N-terminus of the expressed proteins. The strong interaction between histidine hexamers and Ni^{2+} ions permitted purification of the proteins by NI-NTA resin (available commercially from Qiagen,

Figs. 6a and 6b illustrate the behavior of wild-type GFP and the Ser 65→Thr mutant, respectively, upon progressive irradiation with ultraviolet light. Numbers indicate minutes of exposure to illumination at 280 nm; intensity was the same for both samples. Wild-type GFP (Fig. 6a) suffered photoisomerization, as shown by a major change in the shape of the excitation spectrum. Illumination with broad band (240-400 nm) UV caused qualitatively similar behavior but with less increase of amplitude in the 430-500 nm region of the spectrum. The photoisomerization was not reversible upon standing in the dark. This photoisomerization would clearly be undesirable for most uses of wild-type GFP, because the protein rapidly loses brightness when excited at its main peak near 395 nm. The mutant (Fig. 6b) showed no such photoisomerization or spectral shift.

Example 3

GFP cDNAs encoding for Tyr66→His (Y66H), Tyr66→Trp (Y66W), or Ser65→Thr (S65T) were separately further mutagenized by the polymerase chain reaction and transformed into *E. coli* for visual screening of colonies with unusual intensities or colors. Isolation, spectral characterization (Table II and Fig. 7), and –

Random mutagenesis of the gfp cDNA was done by increasing the error rate of the PCR with 0.1 mM $MnCl_2$ and unbalanced nucleotide concentrations. The GFP mutants S65T, Y66H and Y66W had been cloned into the BamH1 site of the expression vector pRSETB (Invitrogen), which includes a T7 promoter and a polyhistidine tag. The GFP coding region (shown in bold) was flanked by the following 5' and 3' sequences: 5'-G GAT CCC CCC GCT GAA TTC **ATG ... AAA TAA** TAA GGA TCC-3'. The 5' primer for the mutagenic PCR was the T7 primer matching the vector sequence; the 3' primer was 5'-GGT AAG CTT TTA TTT GTA TAG TTC ATC CAT GCC-3', specific for the 3' end of GFP, creating a HindIII restriction site next to the stop codon. Amplification was over 25 cycles (1 min at 94° C, 1 min 52° C, 1 min 72° C) using the AmpliTaq polymerase from Perkin Elmer. Four separate reactions were run in which the concentration of a different nucleotide was lowered from 200 μM to 50 μM. The PCR products were

Table II

Fluorescence properties of GFP mutants

Clone	Mutations	Excitation max (nm)	Emission max (nm)	Extinct.Coeff. ($M^{-1}cm^{-1}$)	Quantum yield
P4-3	Y66H Y145F	381	445	14,000	0.38
W7	Y66W N146I M153T V163A N212K	433 (453)	475 (501)	18,000 (17,100)	0.67
W2	Y66W I123V Y145H H148R M153T V163A N212K	432 (453)	480	10,000 (9,600)	0.72
P4-1	S65T M153A K238E	504 (396)	514	14,500 (8,600)	0.54

```
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
 50                  55                  60

TCT TAT GGT GTT CAA TGC TTT TCA AGA TAC CCA GAT CAT ATG AAA CGG      240
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
 65                  70                  75                  80

CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG GAA AGA      288
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                 85                  90                  95

ACT ATA TTT TTC AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT GAA GTC      336
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
            100                 105                 110

AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA AAA GGT ATT      384
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
        115                 120                 125

GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA TTG GAA TAC AAC      432
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
    130                 135                 140

TAT AAC TCA CAC AAT GTA TAC ATC ATG GCA GAC AAA CAA AAG AAT GGA      480
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145                 150                 155                 160

ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT GAA GAT GGA AGC GTT      528
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
                165                 170                 175

CAA CTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC GAT GGC CCT      576
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
            180                 185                 190

GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA CAA TCT GCC CTT TCG      624
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
        195                 200                 205

AAA GAT CCC AAC GAA AAG AGA GAC CAC ATG GTC CTT CTT GAG TTT GTA      672
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
    210                 215                 220

ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA TA           717
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 238 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
 1               5                  10                  15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
            20                  25                  30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
        35                  40                  45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
    50                  55                  60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
```

WHAT IS CLAIMED IS:

1. A composition of matter comprising a fluorescent product derived from a modified form of an *Aequorea* wild-type GFP polypeptide, characterized in that upon oxidation and cyclization of amino acid residues in the modified form corresponding to positions 65 to 67 of wild-type GFP polypeptide sequence [SEQ ID NO:2] a product exhibiting a different excitation and/or emission spectrum from a corresponding product derived from the wild-type GFP polypeptide sequence is formed.

2. A composition according to claim 1, wherein the product exhibits an alteration in the ratio of two main excitation peaks relative to the corresponding product derived from wild-type GFP.

3. A composition according to claim 2, wherein increased fluorescence is exhibited at a shorter-wavelength peak of the two main excitation peaks.

4. A composition according to claim 3, wherein the modified form of the wild-type GFP sequence comprises a replacement of Ser at a position corresponding to position 202 in the wild-type GFP sequence by Phe and a replacement of Thr at a position corresponding to position 203 by Ile.

5. A composition according to claim 2, wherein increased fluorescence is exhibited at a longer-wavelength peak of the two main excitation peaks.

6. A composition according to claim 5, wherein the modified form of the wild-type GFP sequence comprises a replacement of Ile at a position corresponding to position 167 of the wild-type GFP sequence by Val or Thr.

7. A composition according to claim 5, wherein the modified form of the wild-type GFP sequence comprises a replacement of Ser at a position

13. A composition according to claim 1, wherein the product exhibits enhanced emission relative to the corresponding product derived from wild-type GFP.

14. A composition according to claim 13, wherein the modified form of the wild-type GFP sequence comprises a replacement of Ser at a position corresponding to position 65 of the wild-type GFP sequence by an amino acid selected from the group consisting of Ala, Cys, Thr, Leu, Val and Ile.

15. A composition according to claim 14, wherein the amino acid is Cys or Thr.

16. A substantially pure oligonucleotide sequence encoding a modified form of an *Aequorea* wild-type GFP polypeptide sequence according to any one of claims 1-15.

17. A method for monitoring expression of a gene encoding a polypeptide, comprising:

forming a combined sequence comprising the gene and an oligonucleotide sequence according to claim 16, in which combined sequence both the gene and the oligonucleotide sequence are in the same reading frame; and

observing for fluorescence characteristic of a product derived from a polypeptide sequence encoded by the oligonucleotide sequence, the fluorescence indicating expression of a fusion protein encoded by the combined sequence.

22. A method for monitoring transcription of a gene, comprising:

forming a combined sequence comprising upstream regulatory elements of the gene and an oligonucleotide sequence according to claim 16, in which combined sequence both the upstream regulatory elements and the oligonucleotide sequence are in the same reading frame; and

measuring the fluorescence characteristics of a product derived from a polypeptide sequence encoded by the oligonculeotide sequence, the fluorescence indicating transcription of the oligonucleotide under the control of the upstream regulatory elements of the gene.

23. A method for simultaneously monitoring expression of a first gene and a second gene in a single cell, tissue or organism, the first gene encoding a polypeptide different from a polypeptide encoded by the second gene, said method comprising:

forming a first combined sequence comprising upstream regulatory elements of the first gene and a first oligonucleotide sequence according to claim 16, in which first combined sequence both the upstream regulatory elements of the first gene and the first oligonucleotide sequence are in the same reading frame;

forming a second combined sequence comprising the upstream regulatory elements of a second gene and a second oligonucleotide sequence selected from the group consisting of an oligonucleotide sequence encoding a wild-type GFP and an oligonucleotide sequence according to claim 16 different from the first oligonucleotide sequence, in which second combined sequence both the upstream regulatory elements of the second gene and the second oligonucleotide sequence are in the same reading frame; and

observing for fluorescences characteristic of products derived from polypeptide sequences encoded by the first and second oligonucleotide sequences, the fluorescences indicating transcription under the control of the respective upstream regulatory elements of the first and second genes.

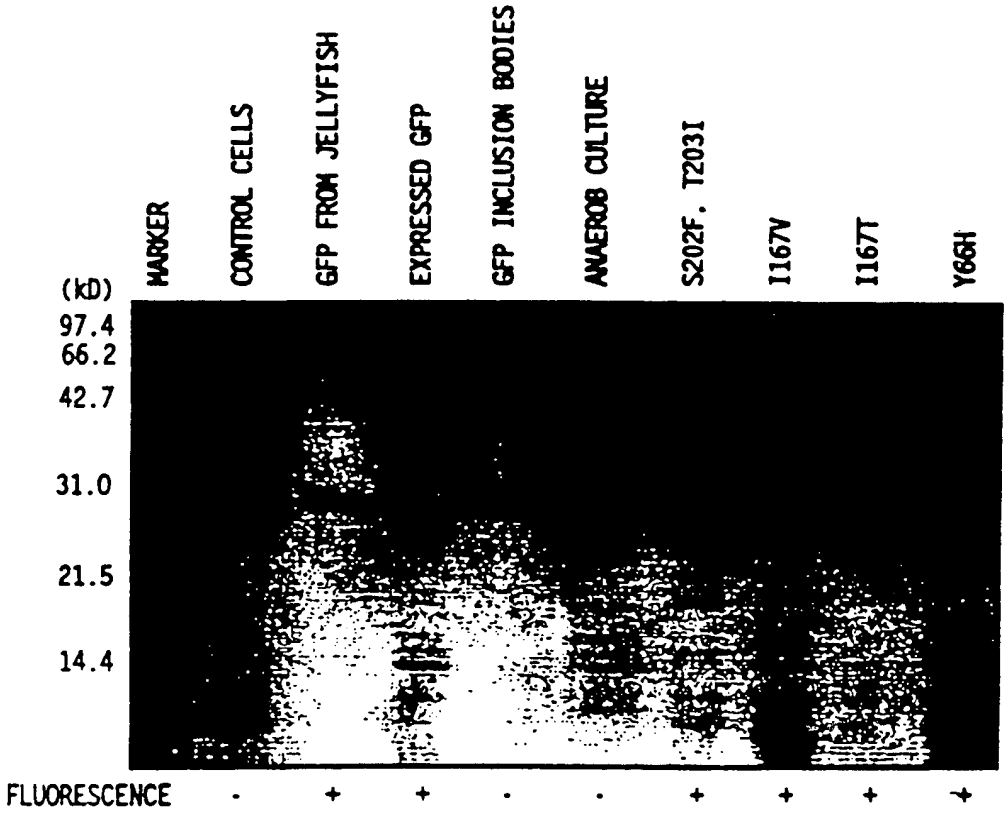

MARKER
CONTROL CELLS
GFP FROM JELLYFISH
EXPRESSED GFP
GFP INCLUSION BODIES
ANAEROB CULTURE
S202F, T203I
I167V
I167T
Y66H
(kD)
97.4
66.2
42.7
31.0
21.5
14.4
FLUORESCENCE - + + - - + + + +

FIG. 1

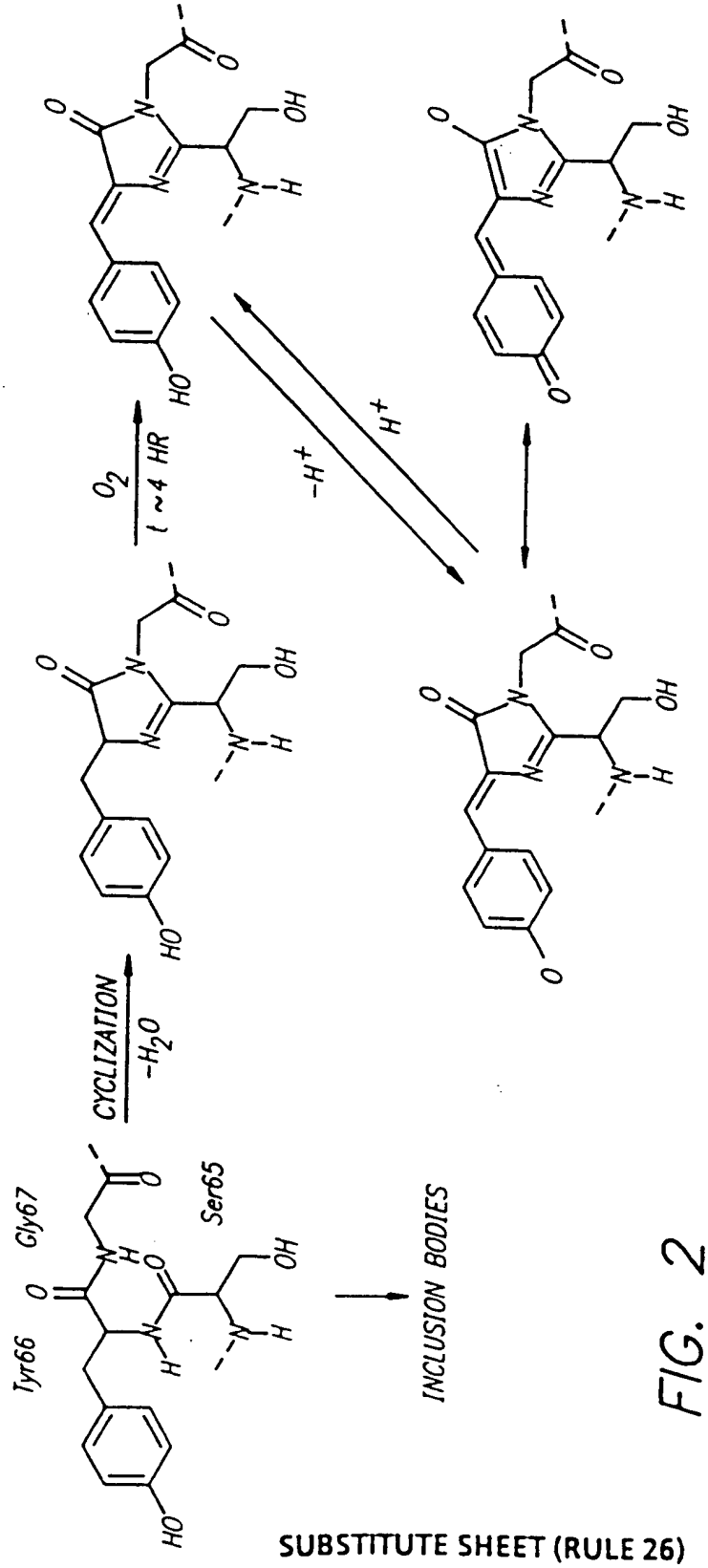

Tyr66
Gly67
Ser65
CYCLIZATION
$-H_2O$
O_2
t ~4 HR
$-H^+$
H^+
INCLUSION BODIES

FIG. 2

FIG. 3a

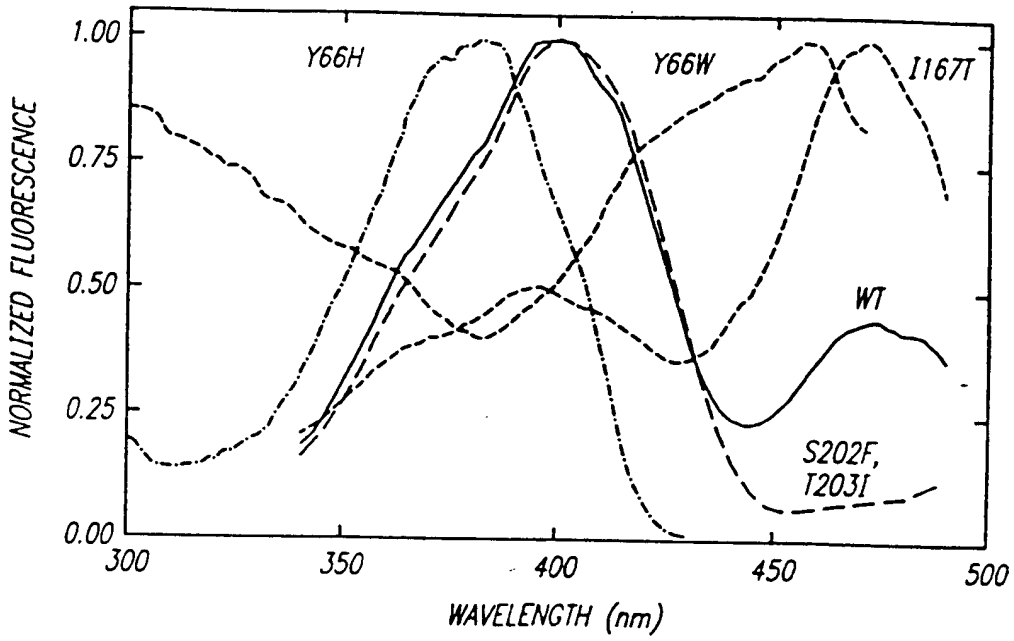

1.00
0.75
0.50
0.25
0.00
NORMALIZED FLUORESCENCE
300
350
400
450
500
WAVELENGTH (nm)
Y66H
Y66W
I167T
WT
S202F,
T203I

FIG. 3b

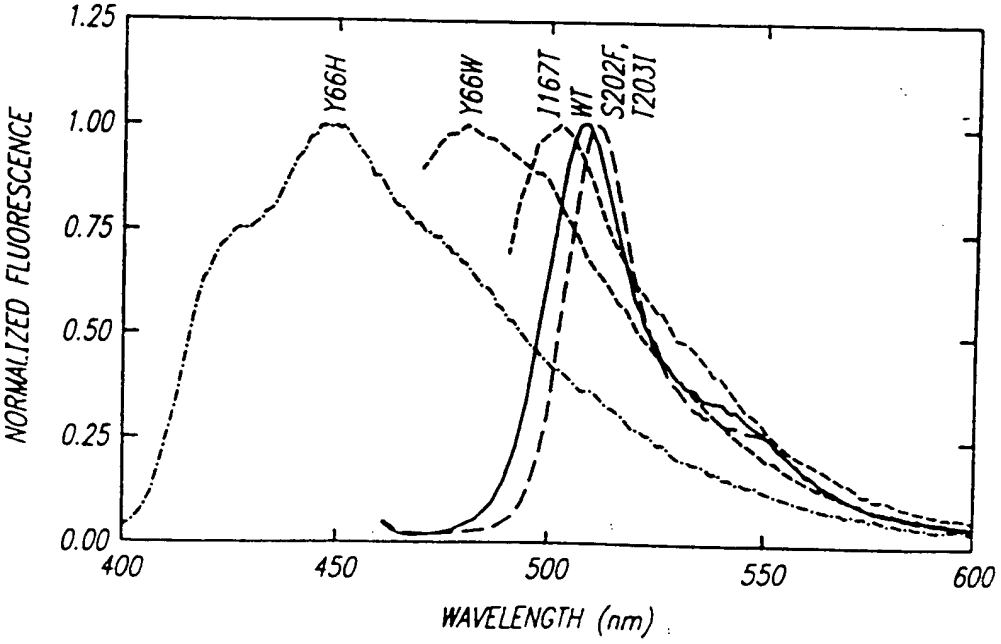

1.25
1.00
0.75
0.50
0.25
0.00
NORMALIZED FLUORESCENCE
400
450
500
550
600
WAVELENGTH (nm)
Y66H
Y66W
I167T
WT
S202F,
T203I

FIG. 4a

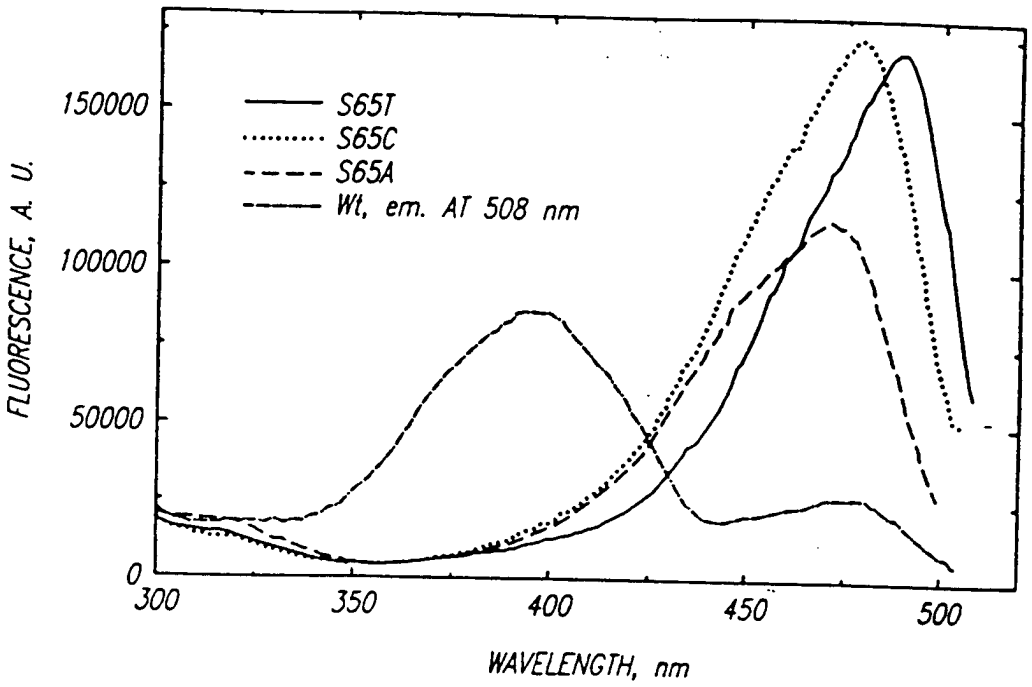

S65T
S65C
S65A
Wt, em. AT 508 nm
150000
100000
50000
0
300
350
400
450
500
FLUORESCENCE, A. U.
WAVELENGTH, nm

FIG. 4b

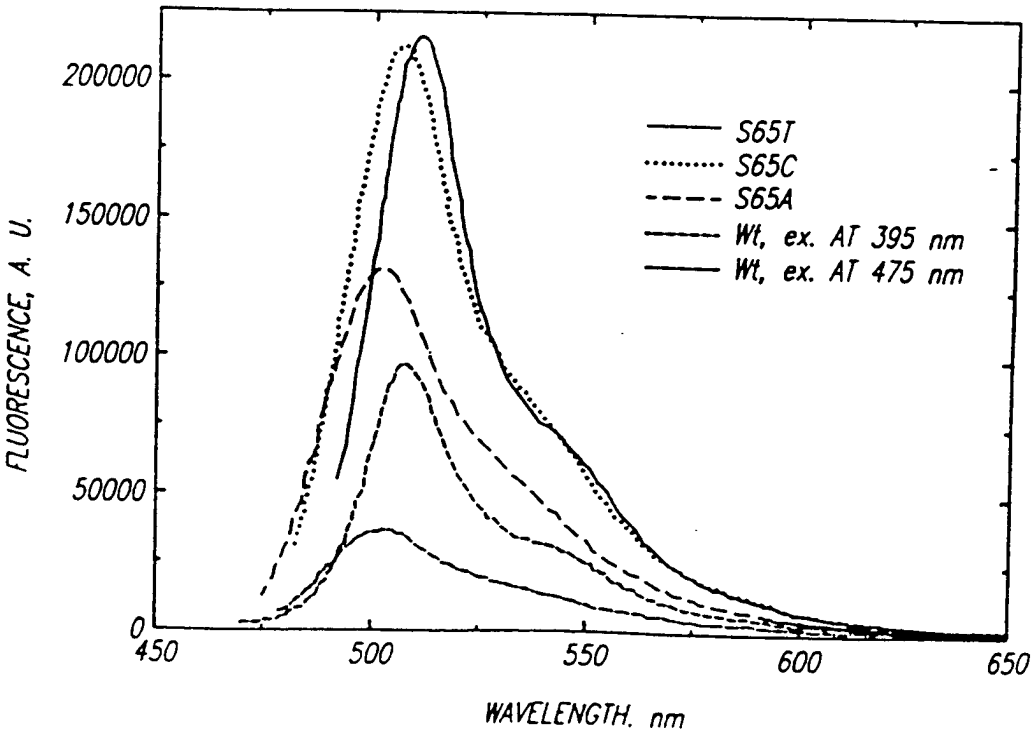

S65T
S65C
S65A
Wt, ex. AT 395 nm
Wt, ex. AT 475 nm
200000
150000
100000
50000
0
450
500
550
600
650
FLUORESCENCE, A. U.
WAVELENGTH. nm

FIG. 5

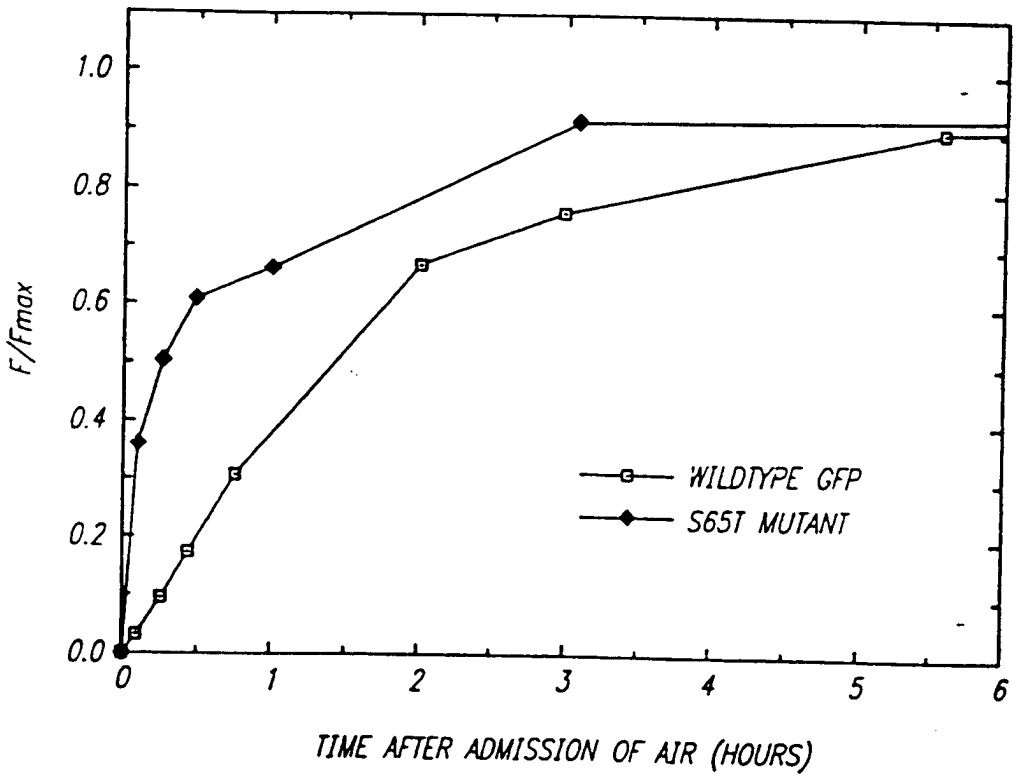

1.0
0.8
0.6
0.4
0.2
0.0
F/Fmax
0
1
2
3
4
5
6
WILDTYPE GFP
S65T MUTANT
TIME AFTER ADMISSION OF AIR (HOURS)

FIG. 6a

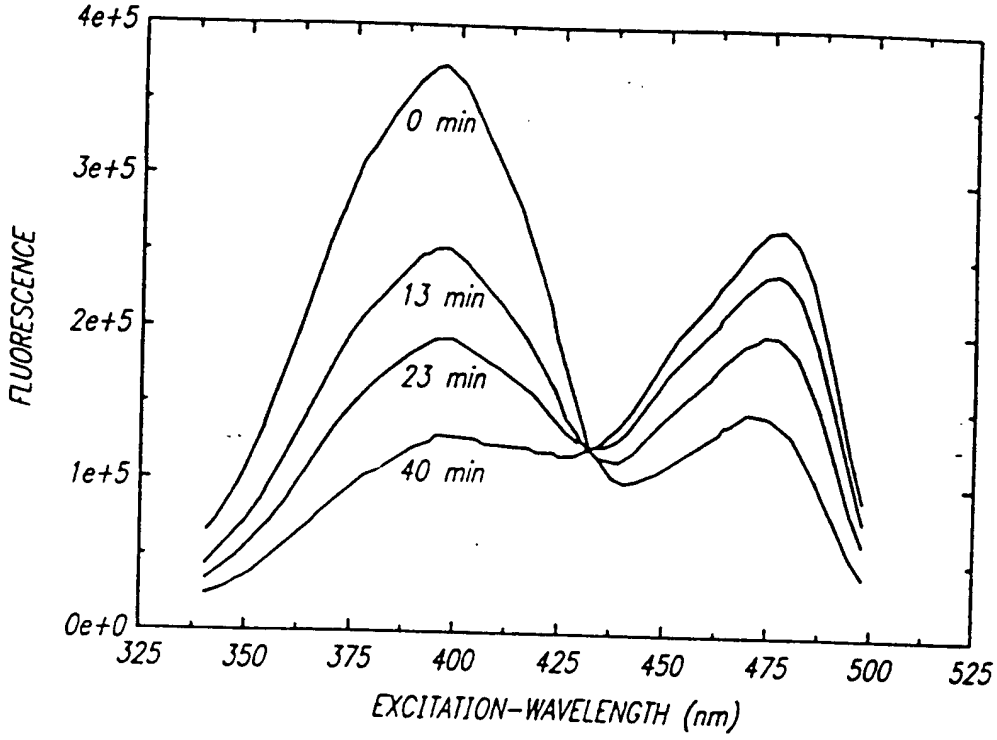
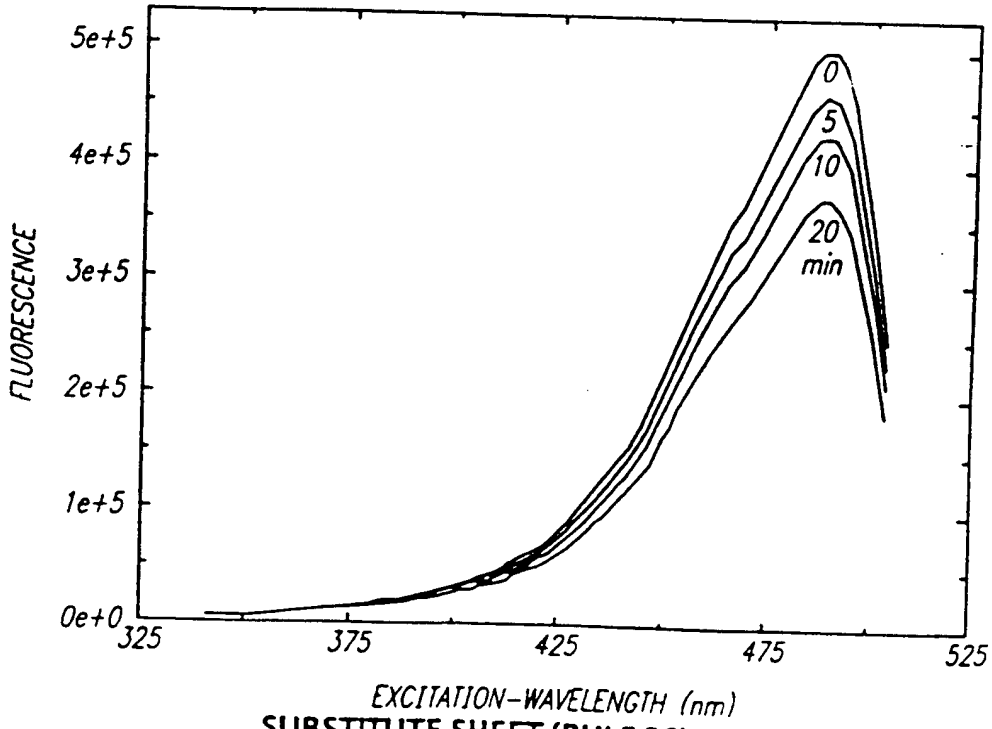

4e+5
3e+5
2e+5
1e+5
0e+0
FLUORESCENCE
0 min
13 min
23 min
40 min
325
350
375
400
425
450
475
500
525
EXCITATION-WAVELENGTH (nm)

FIG. 6b


5e+5
4e+5
3e+5
2e+5
1e+5
0e+0
FLUORESCENCE
0
5
10
20
min
325
375
425
475
525
EXCITATION-WAVELENGTH (nm)

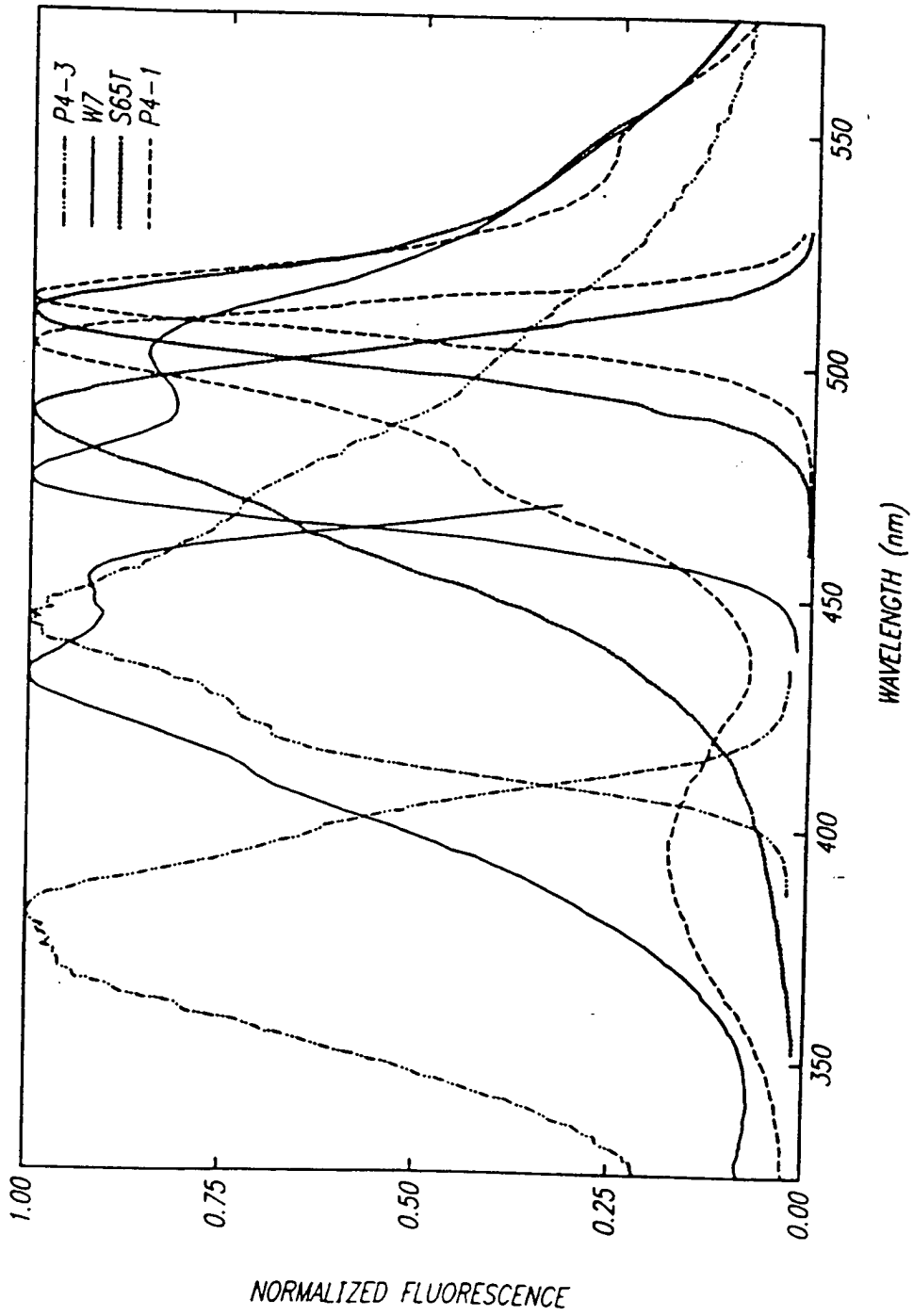

P4-3
W7
S65T
P4-1
1.00
0.75
0.50
0.25
0.00
350
400
450
500
550
NORMALIZED FLUORESCENCE
WAVELENGTH (nm)

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/14692

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :435/6; 530/350; 536/23.5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 530/350; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, CAPLUS, SCISEARCH
search terms: green fluorescent protein

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X ------- P, Y	FEBS Letters, Vol. 367, No. 2, issued 26 June 1995, Ehrig et al., "Green fluorescent protein mutants with altered fluorescence excitation spectra", pages 163-166, see pages 163-165.	1-3, 5, 8, 16-18, 20, 22, 23 ---------- 4, 6, 7, 9-15, 19, 21
P, X ------- P, Y	Proceedings of the National Academy of Sciences, Vol. 91, No. 26, issued 20 December 1994, Heim et al., "Wavelength mutations and posttranslational autoxidation of green fluorescent protein", pages 12501-12504, see entire document.	1-6, 8, 9, 13, 16-18, 20, 22, 23 ---------- 7, 10, 14, 15, 19, 21

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 JANUARY 1996	21 FEB 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer SCOTT D. PRIEBE Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/14692

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07H 21/04; C07K 14/435; C12Q 1/25, 1/68